CLAIMS

1   1.   A computer programmed to undertake method acts for querying for data using a

2   query, the method acts undertaken by the computer including:

3   for at least some data vectors in a data space, generating respective approximations

4   in polar coordinates; and

5   based on the approximations, returning "k" nearest neighbors to the query.

1   2.   The computer of Claim 1, wherein the method acts further comprise:

2   dividing the data space into plural cells; and

3   representing at least one data point in at least one cell in polar coordinates with

4   respect to the at least one cell.

1   3.   The computer of Claim 2, wherein the data space has "d" dimensions and the

2   method acts further comprise:

3   determining a number of "b" bits to be assigned to each cell; and

4   dividing the data space into 2^{bd} cells.

1   4.   The computer of Claim 1, wherein each approximation defines a lower bound d_{min},

2   and the method acts further comprise:

3   generating a candidate set of approximations based at least on the lower bounds

4   d_{min} of the approximations.

1   5.   The computer of Claim 4, wherein the query can be represented by a query vector

2 q, and the method acts further comprise:

3   adding a first approximation having a first lower bound d_{min1} to the candidate set

4 if $d_{min1} < \text{k-NN}^{dist}(q)$, wherein $\text{k-NN}^{dist}(q)$ is the kth largest distance between the query

5 vector q and nearest neighbor vectors p.

1   6.   The computer of Claim 5, wherein the method acts further comprise using the

2 candidate set to return "k" nearest neighbors vectors p to the query vector q.

1   7.   The computer of Claim 6, wherein not all vectors p corresponding to

2 approximations in the candidate set are examined to return the "k" nearest neighbors.

1   8.   A computer program product including a program of instructions having:

2   computer readable code means for generating approximations including local polar

3 coordinates of at least some data vectors p in at least one data set having a dimensionality

4 of "d", the local polar coordinates being independent of "d"; and

5   computer readable code means for using the approximations to return "k" nearest

6 neighbors to a query.

1   9.   The computer program product of Claim 8, wherein the means for generating

2   generates respective approximations of data vectors **p** in local polar coordinates.

1   10.   The computer program product of Claim 9, further comprising:

2   computer readable code means for dividing the data space into plural cells; and

3   computer readable code means for representing each approximation in polar

4   coordinates with respect to one of the cells.

1   11.   The computer program product of Claim 10, wherein the data space has "d"

2   dimensions, further comprising:

3   computer readable code means for determining a number of "b" bits to be assigned

4   to each cell; and

5   computer readable code means for dividing the data space into 2^{bd} cells.

1   12.   The computer program product of Claim 9, wherein each approximation defines

2   a lower bound d_{min} and an upper bound d_{max}, and the product further comprises:

3   computer readable code means for generating a candidate set of approximations

4   based at least on the lower bounds d_{min} and upper bounds d_{max} of the approximations.

1   13.   The computer program product of Claim 12, further comprising:

2   computer readable code means for adding a first approximation having a first

3   lower bound d_{min1} to the candidate set if $d_{min1} < k\text{-}NN^{dist}$ (**q**), wherein $k\text{-}NN^{dist}$ (**q**) is the

4   k^{th} largest distance between the query vector **q** and nearest neighbor vectors **p** associated

5   with approximations in the candidate set.

1   14.   The computer program product of Claim 13, further comprising computer readable

2   code means for using the candidate set to return "k" nearest neighbors vectors **p** to the query

3   vector **q**.

1   15.   A computer-implemented method for finding, in a data space, "k" closest data

2   vectors **p** to a query vector **q**, comprising:

3   rendering approximations of at least some of the data vectors **p** using local polar

4   coordinates;

5   filtering the approximations; and

6   after filtering, returning the "k" closest data vectors **p**.

1   16.   The method of Claim 15, further comprising:

2   dividing the data space into plural cells; and

3   representing each approximation in polar coordinates with respect to one of the

4   cells.

1   17.   The method of Claim 16, wherein the data space has "d" dimensions and the

2   method further comprises:

3   determining a number of "b" bits to be assigned to each cell; and

4   dividing the data space into 2^{bd} cells.

1   18.   The method of Claim 15, wherein each approximation defines a lower bound d_{min},

2   and the method further comprises:

3   generating a candidate set of approximations based at least on the lower bounds

4   d_{min} of the approximations.

1   19.   The method of Claim 18, further comprising:

2   adding a first approximation having a first lower bound d_{min1} to the candidate set

3   if $d_{min1} < $ k-NN^{dist} (q), wherein k-NN^{dist} (q) is the kth largest distance between the query

4   vector q and nearest neighbor vectors p associated with approximations in the candidate

5   set.

1   20.   The method of Claim 19, further comprising using the candidate set to return "k"

2   nearest neighbors vectors p to the query vector q.

1   21.   The method of Claim 20, wherein not all data vectors p corresponding to

2   approximations in the candidate set are examined to return the "k" nearest neighbors vectors p.

1   22.   The computer of Claim 4, wherein each approximation defines an upper bound

2   d_{max}, and the method acts further comprise:

3   generating a candidate set of approximations based at least on the upper bounds

4   d_{max} of the approximations.

1   23.   The computer program product of Claim 12, wherein each approximation defines

2   an upper bound d_{max}, and the product further comprises:

3   computer readable code means for generating a candidate set of approximations

4   based at least on the upper bounds d_{max} of the approximations.

1   24.   The computer of Claim 1, wherein each approximation defines an upper bound

2   d_{max}, and the method acts further comprise:

3   generating a candidate set of approximations based at least on the upper bounds

4   d_{max} of the approximations.